

## **Metodologia para Estimaco Robusta com Varincias Residuais Heterogneas em Populaes Multirraciais**



*Empresa Brasileira de Pesquisa Agropecuária  
Embrapa Pecuária Sul  
Ministério da Agricultura, Pecuária e Abastecimento*

# ***Documentos***118

## **Metodologia para Estimação Robusta com Variâncias Residuais Heterogêneas em Populações Multirraciais**

*Fernando Flores Cardoso  
Maurício Morgado de Oliveira*

Exemplares desta publicação podem ser adquiridos na:

**Embrapa Pecuária Sul**

BR 153, Km 633. Caixa postal 242

96401-970 - Bagé – RS

Fax: 55 53 3240-4650

<https://www.embrapa.br/fale-conosco/sac>

**Comitê Local de Publicações**

Presidente: *Claudia Cristina Gulias Gomes*

Secretária-Executiva: *Graciela Olivella Oliveira*

Membros: *Estefanía Damboriarena, Fernando Flores Cardoso, Jorge Luiz Sant'Anna dos Santos, Lisiane Bassols Brisolara, Marco Antônio Karam Lucas, Naylor Bastiani Perez, Renata Wolf Suñé*

Supervisor editorial: *Lisiane Bassols Brisolara*

Revisor de texto: *Comitê Local de Publicação*

Normalização bibliográfica: *Graciela Olivella Oliveira*

Editoração eletrônica: *Núcleo de Comunicação Organizacional*

Foto da capa: *Gabriel Becco*

**1ª edição**

**Todos os direitos reservados**

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei N° 9.610).

**Dados Internacionais de Catalogação na Publicação (CIP)**

Embrapa Pecuária Sul

---

Cardoso, Fernando Flores

Metodologia para estimação robusta com variâncias residuais heterogêneas em populações multirraciais [recurso eletrônico] / Fernando Flores Cardoso, Maurício Morgado de Oliveira. - Bagé : Embrapa Pecuária Sul, 2015.

(Documentos / Embrapa Pecuária Sul, ISSN 1982-5390 ; 118)

Sistema requerido: Adobe Acrobat Reader

Modo de acesso: World Wide Web: <[www.embrapa.br](http://www.embrapa.br)>

Título da página Web (acesso em 30 nov. 2015)

1. Genética quantitativa. 2. Melhoramento genético animal. 3. Gado de corte. I. Oliveira, Maurício Morgado de. II. Série.

CDD 636.0821

# **Autores**

## **Fernando Flores Cardoso**

Médico Veterinário, Doutor (Ph.D.) em Bioinformática  
ênfase em Estatística Genômica, pesquisador da  
Embrapa Pecuária Sul,  
Caixa Postal 242, BR 153 Km 633,  
CEP 96401-970 - Bagé, RS – Brasil

## **Maurício Morgado de Oliveira**

Zootecnista, Doutor em Zootecnia - ênfase em  
Melhoramento Animal,  
Pós-doutorando da Universidade Federal de Santa  
Maria.

# **Apresentação**

O investimento em pesquisa, desenvolvimento e inovação (PD&I) é fundamental para o progresso socioeconômico e a independência de uma nação. Dentre os temas prioritários para a geração de produtos e tecnologias através da pesquisa científica na Embrapa Pecuária Sul está o melhoramento genético animal. A disseminação dos conhecimentos gerados nessa área possibilita que os beneficiários do processo, ou seja, produtores, técnicos, estudantes e demais interessados nas novas tecnologias genéticas, possam selecionar e acasalar animais mais produtivos e eficientes, que venham a melhorar o resultado econômico nas cadeias produtivas.

Neste contexto, a Embrapa Pecuária Sul utiliza as publicações da Série Embrapa como uma das ferramentas formais estratégicas de transferência de tecnologia direcionadas às cadeias produtivas da carne bovina e ovina, do leite e da lã para a região Sul do Brasil.

A presente publicação é mais um exemplo deste esforço institucional. Nesta obra é apresentada uma metodologia científica que aumenta a acurácia do cálculo de valores genéticos de animais oriundos do cruzamento entre raças e criados em condições extensivas, típicas dos Campos Sul-brasileiros.

Esperamos que o emprego dos conhecimentos disponibilizados nesta obra nos programas de avaliação genética de populações multirraciais possa colaborar com o aumento do progresso genético dos rebanhos e que a genética comprovadamente superior se dissemine para permear suas vantagens na atividade pecuária do sul do Brasil.

*Fernando Flores Cardoso*  
Chefe de PD&I

# Sumário

<b>Introdução</b>	07
<b>Material e Métodos</b>	09
Conjunto de Dados Nelore-Hereford	09
Especificação dos Modelos	09
Definição dos Parâmetros Genéticos	12
Critério de Escolha de Modelos	12
<b>Resultados e Discussão</b>	13
Escolha do Modelo	13
Avaliação das Fontes de Heteroscedasticidade	14
Deteção e Mitigação de Dados Extremos ("Outliers")	16
Herdabilidades	20
Avaliações Genéticas	22
Comparação dentro de Composições Raciais	22
Comparação entre Composições Raciais	24
<b>Conclusões</b>	28
<b>Referências</b>	30

# Metodologia para Estimação Robusta com Variâncias Residuais Heterogêneas em Populações Multirraciais

---

*Fernando Flores Cardoso  
Maurício Morgado de Oliveira*

## Introdução

No melhoramento genético, o progresso em desempenho das populações devido a seleção depende fundamentalmente do intervalo entre gerações, da intensidade de seleção, da variabilidade genética disponível e da precisão dos valores genéticos estimados.

As avaliações genéticas são utilizadas atualmente para combinar os dados de diversos rebanhos e obter estimativas diretamente comparáveis dos valores genéticos para todos os animais na população e a precisão dessas estimativas é uma função da qualidade da informação fenotípica e de pedigree disponível.

Neste contexto, o desempenho de bovinos de corte é geralmente avaliado através de sistemas e de ambientes altamente diversos de produção, com qualidade de dados muitas vezes comprometida por ocorrências de erros de registro, de tratamento preferencial e do efeito potencial de doenças ou lesões. Esses fatores podem afetar substancialmente o desempenho dos animais, mas não são considerados explicitamente no modelo estatístico e poderão gerar dados extremos ("outliers"), enviesando de forma significativa as estimativas dos parâmetros genéticos e as avaliações genéticas (STRANDEN; GIANOLA, 1998).

A maioria dos programas de avaliação genética tem lidado com esta situação através da edição de dados, eliminando observações que são consideradas extremamente afastadas da média fenotípica do grupo de contemporâneos (GC) ao qual o animal pertence (usualmente três ou mais desvios padrão) ou se a relação observação/média de seu GC se encontra fora do intervalo 60% a 140% (BERTRAND; WIGGANS, 1998). Apesar de que estas estratégias possam parecer adequadas para erros de coleta de informação muito óbvios (p. ex., erros de digitação), podem ser ineficientes se a verdadeira distribuição dos dados se afasta da normalidade, estando naturalmente sujeita a ocorrência de observações extremas ou se as variâncias das observações são heterogêneas entre os

Os modelos hierárquicos de Bayes proporcionam uma estrutura geral para lidar de forma mais apropriada com dados extremos e caracterizados por diferentes tipos de heterogeneidade através dos vários ambientes (heteroscedasticidade). Modelos robustos com resíduos heteroscedásticos propostos para avaliações genéticas por Cardoso et al. (2005) combinam características de modelos estruturais de variância (FOULLEY et al., 1992; KIZILKAYA; TEMPELMAN, 2005) com distribuições de caudas longas, tais como a *t* de Student ou a Slash (LANGE; SINSHEIMER, 1993; ROSA et al., 2003), sendo uma forma alternativa de mitigar o efeito de dados extremos sem a necessidade de eliminação arbitrária de observações.

Como será demonstrado, as especificações deste modelo podem ser sobrepostas em um modelo hierárquico que explique variações genéticas heterogêneas através de grupos raciais em avaliações genéticas multirraciais (CARDOSO; TEMPELMAN, 2004).

Os objetivos deste documento são: 1) Apresentar modelos para avaliação genética de populações multirraciais na presença de dados extremos e heterogeneidade de variância ambiental; 2) Comparar esses modelos com o modelo clássico baseado na pressuposição de distribuição normal ou gaussiana dos dados; 3) Aplicar os modelos propostos em conjunto de dados de ganho pós-desmame de Nelore-Hereford criados em ambientes diversos, para caracterizar a magnitude de fontes potencialmente



importantes de heterogeneidade de variância e a decorrente influência na herdabilidade específica de cada grupo racial; 4) Demonstrar a utilidade dos modelos bayesianos heteroscedásticos robustos para a detecção e mitigação de dados extremos; e 5) Avaliar as consequências da aplicação desses modelos em avaliações genéticas multirraciais.

## **Material e Métodos**

### **Conjunto de dados Nelore-Hereford**

Os dados utilizados nas análises foram 22.717 registros de ganho pós-desmama (GPD) (40.082 animais nos arquivos de pedigree) de uma população composta de animais Hereford, Nelore e seus cruzamentos, pertencentes a 15 rebanhos incluídos no programa de melhoramento da “Conexão Delta G” no Brasil. Os animais foram criados em condições de pastejo em três regiões diferentes: Região 1 entre 14°S e 16°S (5.410 registros); Região 2 entre 21°S e 23°S (3.110 registros), e a Região 3 entre 30°S e 32°S (14.197 registros). A distribuição dos grupos raciais foi altamente desbalanceada entre regiões, com predominância de progênies F1 de vacas Nelore nas regiões tropicais 1 e 2, e Hereford puros na região sub-tropical 3. O GPD médio  $\pm$  erro padrão foi  $98,2 \pm 41,2$  Kg. Maiores detalhes da estrutura dos dados podem ser encontrados em Cardoso e Tempelman (2004).

### **Especificação dos modelos**

Os modelos propostos e utilizados combinam heteroscedasticidade, ou seja, variâncias residuais heterogêneas para subclasses de GC, sexo, proporção racial e heterozigosidade, e robustez a dados extremos. Uma representação fatorial  $3 \times 2$  foi utilizada para a especificação de seis modelos diferentes de variância dos resíduos, com o fator 1 representando a distribuição (Gaussiana,  $t$  de Student ou Slash) e o segundo fator especificando a variabilidade (homoscedástica vs heteroscedástica) (CARDOSO et al., 2005). Os seis modelos de variância

dos resíduos resultantes são os seguintes: 1) Gaussiano homoscedástico (G-HO); 2) t Student homoscedástico (T-HO); 3) Slash homoscedástico (S-HO); 4) Gaussiano heteroscedástico (G-HE); 5) t de Student heteroscedástico (T-HE); e 6) Slash heteroscedástico (S-HE).

Para os seis modelos de variância residual foi especificado o mesmo modelo misto como uma função aditiva dos seguintes parâmetros:

$$y_{ijklmn} = \mu + reg_i + gc_j + sex_k + \beta_{PPD} (ppd_l) + \beta_{IV} (iv_m) + \beta_{IV^2} (iv_m^2) \\ + \beta_{sex \times A} (sex_k \times f_n) + \beta_{PPD \times A} (ppd_l \times f_n) + \beta_{IV \times A} (iv_m \times f_d) + \beta_{IV^2 \times A} (iv_m^2 \times f_d) \\ + \beta_A (f_n) + \beta_D (\delta_n) + \beta_{AA} (2[1 - f_n] f_n) + u_n + e_{ijklmn}$$

onde  $y_{ijklmn}$  é a observação de desempenho do  $n^{ésimo}$  animal da  $i^{ésima}$  região, do  $j^{ésimo}$  grupo contemporâneo (GC) e do  $k^{ésimo}$  sexo;  $\mu$  é a media geral;  $reg_i$  é o efeito da  $i^{ésima}$  região ( $i = 1, 2, 3$ );  $gc_j$  é o efeito do  $j^{ésimo}$  GC ( $j = 1, 2, \dots, 940$ ), com  $gc_j \sim N(0, \sigma_{gc}^2)$  para todo  $j$ ;  $sex_k$  é o efeito do  $k^{ésimo}$  sexo ( $k = 1$  ou  $k = 0$  para macho ou fêmea respectivamente);  $\beta$ 's são os coeficientes de regressão associados com o período de avaliação pós-desmame, em dias (PPD,  $106 \leq PPD \leq 483$ ), idade da vaca em anos (IV,  $2 \leq IV \leq 12$ ), efeitos "fixos" aditivos (A) para a raça Nelore, efeito de dominância (D) Nelore-Hereford, efeitos epistáticos aditivos por aditivos (AA) entre Nelore-Hereford, e as interações entre A e sexo, PPD e IV. Os coeficientes  $f_n$  ( $f_d$ ) representam a proporção esperada de alelos Nelore por animal  $n$  (vaca  $d$ ), enquanto  $\delta_n$  é o coeficiente de heterozigose, obtido como a probabilidade de um dos alelos ser derivado da raça Nelore e o outro alelo ser derivado da raça Hereford para um locus escolhido aleatoriamente no indivíduo  $n$ . Já coeficiente de interações epistáticas aditivas  $\times$  aditivas para o animal  $n$ ,  $2[1 - f_n] f_n$ , foi obtido com base na definição de perdas de recombinação de Kinghorn (1980). Além disso,  $u_n$  é o efeito genético aditivo do  $n^{ésimo}$  animal, assumindo que  $\mathbf{u} = \{u_n\}_{n=1}^{40082} \sim N(\mathbf{0}, \mathbf{G})$ , onde  $\mathbf{G}$  é a matriz de (co)variâncias genéticas multirraciais que é uma função das variâncias genéticas raciais específicas para as raças Nelore e Hereford da variância de segregação entre ambas as raças e das relações genéticas (parentesco) entre os animais (LO et al., 1993).

Finalmente,  $e_{ijklmn}$  é o termo do resíduo com distribuição de acordo às seis especificações alternativas apresentadas acima. Mais especificamente,  $e_{ijklmn} \sim N(0, \sigma_{ijklmn}^2)$  onde  $\sigma_{ijklmn}^2 = \sigma_e^2 \gamma_A \tau_k \gamma_D w_n^{-1}$  é a variância residual específica para a observação do animal  $n$ , determinada através de uma função multiplicativa da variância residual global ( $\sigma_e^2$ ) e fatores escalares para GC ( $\gamma_A$ ) sexo ( $\tau_k$ ), raça ( $\gamma_D$ ) e heterozigosidade ( $\gamma_D$ ) para os três modelos de erros heteroscedásticos (G-HE, T-HE, e S-HE). Por exemplo, a variância residual para um determinado macho  $F_1$  (isto é,  $f_1^j = 0.50$ ,  $f_{12}^j = 1.00$ ) com registro no GC 2 baseado no modelo de erro T-HE é determinado por  $\sigma_e^2 \gamma_2 \tau_1 (\gamma_1)^{0.50} (\gamma_2)^{1.00} w_n^{-1}$ .

Para os três modelos do erro homoscedástico (G-HO, T-HO, e S-HO), a variância do erro é especificada de forma mais simples para todas as observações como  $\sigma_{ijklmn}^2 = \sigma_e^2 w_n^{-1}$ . Aqui  $w_n$  é uma variável de ponderação para a observação do animal  $n$ . A distribuição de  $w_n$  define a robustez do modelo a dados extremos. Por exemplo, as especificações dos erros na distribuição Gaussiana (G-HO e G-HE) são obtidas considerando  $w_n = 1$  para todo  $n$  e esses modelos não apresentam nenhuma proteção contra o efeito de dados extremos. Por outro lado, as especificações de distribuições de caudas longas sobre os valores de  $w_n$ , conferem ao modelo robustez a dados extremos que se desviem de uma distribuição Normal ou Gaussiana (LANGE; SINSHEIMER, 1993). Foram consideradas especificamente as distribuições  $t$  de Student e Slash, mas outras alternativas são discutidas em Lange e Sinsheimer (1993) e em Rosa et al. (2003). A distribuição residual  $t$  de Student (T-HE e T-HO) foi obtida assumindo que  $w_n \sim \text{Gamma}\left(\frac{v_e}{2}, \frac{v_e}{2}\right)$  para todo  $n$  com  $w_n > 0$ , onde  $v_e > 0$  representa os graus de liberdade ou parâmetro de robustez dessa distribuição. A distribuição residual Slash (S-HE e S-HO) foi especificada alternativamente por  $w_n \sim p(w_n | v_e = v_e w_n^{-1})$  para todo  $n$ , com  $0 < w_n \leq 1$  e  $v_e > 0$ . Detalhes adicionais nessas especificações sob o enfoque bayesiano, bem com densidades a priori e posteriores condicionais completas podem ser encontrados em Cardoso et al. (2005).

A inferência Bayesiana foi baseada em métodos Monte Carlo via Cadeias de Markov (MCMC) de 200.000 ciclos, após um período de descarte de 15.000 ciclos para os seis modelos. Médias, modas, intervalos de probabilidade e erros padrão a posteriori dos parâmetros obtidas a partir das amostras das cadeias MCMC foram utilizadas para inferência.

Correlações de ordem de Kendall entre a média a posteriori dos efeitos genéticos aditivos obtidos pelos diferentes modelos foram utilizados para comparar o ordenamento dos animais para GPD na avaliação genética para realizar seleção dentro de genótipo.

## Definição dos Parâmetros Genéticos

A variância genética aditiva  $\sigma_{A_g}^2$  de qualquer indivíduo pertencente ao grupo racial  $g$  foi obtida por  $\sigma_{A_g}^2 = f_1^g \sigma_{A_1}^2 + f_2^g \sigma_{A_2}^2 + 2(f_1^s f_2^s + f_1^d f_2^d) \sigma_{S_{12}}^2$  (LO et al., 1993), onde  $f_b^i$  indica a proporção alélica da raça  $b$  no grupo racial  $i$ . Aqui  $b = 1$  e  $b = 2$  especificam Nelore e Hereford, respectivamente, e o sobrescrito  $i$  especifica o grupo racial próprio do indivíduo ( $i = g$ ), o grupo racial do pai desse indivíduo ( $i = s$ ) ou o grupo racial da mãe do indivíduo ( $i = d$ ). Dentro de cada um dos três modelos de erro homoscedástico (G-HO, T-HO, e S-HO), a variância residual marginal ( $\sigma_E^2$ ) é o mesmo para cada grupo racial, sendo  $\sigma_e^2$ ,  $\frac{v_e}{v_e - 2} \sigma_e^2$ , e  $\frac{v_e}{v_e - 1} \sigma_e^2$ , respectivamente. Para o modelo G-HE, a variância residual marginal específica para o grupo racial  $g$  ( $\sigma_{E_g}^2$ ) foi determinada usando a expressão  $\sigma_{E_g}^2 = \sigma_e^2 \tau_1^s \gamma_1^g f_1^g f_2^g$ , notando que  $E(v_j) = 1$  em todos os GC. Similarmente, a variância residual marginal foi especificada sendo  $\sigma_{E_g}^2 = \frac{v_e}{v_e - 2} \sigma_e^2 \tau_1^s \gamma_1^g f_1^g f_2^g$  para T-HE; e por  $\sigma_{E_g}^2 = \frac{v_e}{v_e - 1} \sigma_e^2 \tau_1^s \gamma_1^g f_1^g f_2^g$  para S-HE. Aqui,  $\bar{x}_1 = 0.3197$  representa a proporção dos terneiros com registros que eram machos. Alternativamente, ajustando  $\bar{x}_1 = 1$  especifica-se  $\sigma_{E_g}^2$  em relação aos terneiros machos e  $\bar{x}_1 = 0$  especifica-se  $\sigma_{E_g}^2$  relativo às terneiras fêmeas dentro do grupo racial  $g$ . A variância fenotípica marginal para animais no grupo racial  $g$  foi então especificada como  $\sigma_{P_g}^2 = \sigma_{E_g}^2 + \sigma_{A_g}^2$ , tal que a herdabilidade aditiva foi  $h_{A_g}^2 = \sigma_{A_g}^2 / \sigma_{P_g}^2$ .

## Critério de escolha de modelos

O pseudo fator de Bayes (PBF) foi utilizado como meio para escolha do modelo de melhor ajuste aos dados. O PBF é calculado pela avaliação da densidade do primeiro estágio na Equação [1] para cada ciclo da cadeia MCMC, onde  $p(y_f | y_{(-f)} M_f)$  é a ordenada preditiva condicional (CPO) para a

observação  $y_j$  sob o modelo  $M_f$ ,  $f = \text{G-HO, T-HO, ..., S-HE}$ . A CPO é uma densidade de validação cruzada, sugerindo que valores de  $y_j$  são mais prováveis quando o modelo  $M_f$  é ajustado a todas as demais observações (denotadas por  $\mathbf{y}_{(-j)}$ ) exceto  $y_j$ . Se  $\mathfrak{G} = [\boldsymbol{\beta}' \quad \mathbf{u}' \quad \sigma_e^2 \quad \boldsymbol{\tau}' \quad \mathbf{v}' \quad \mathbf{w}']$  é o conjunto de todos os parâmetros no modelo, uma aproximação de MCMC para a CPO é obtida pela seguinte média harmônica:  $p(y_j | \mathbf{y}_{(-j)}, M_f) \approx \left( \frac{1}{G} \sum_{i=1}^G p^{-1}(y_j | \mathfrak{G}^{(i)}, M_f) \right)^{-1}$  com  $\mathfrak{G}^{(i)}$  denotando uma amostra de  $\mathfrak{G}$  após a convergência da cadeia MCMC, com  $i = 1, 2, \dots, G = 20.000$ .

A partir dessas CPO podemos calcular o logaritmo da verossimilhança marginal (LVM) como uma medida de ajuste global dos modelos a todas as observações:  $LVM_f = \sum_{j=1}^n \log p(y_j | \mathbf{y}_{(-j)}, M_f) \approx \sum_{j=1}^n \left( \frac{1}{G} \sum_{i=1}^G p^{-1}(y_j | \mathfrak{G}^{(i)}, M_f) \right)^{-1}$ . Quanto maior o valor de LVM melhor o ajuste do modelo. Finalmente, para comparar, por exemplo, os modelos  $M_{\text{G-HO}}$  e  $M_{\text{T-HO}}$ , o PBF correspondente é determinado ser  $PBF_{\text{G-HO: T-HO}} = \exp(LML_{\text{G-HO}} - LML_{\text{T-HO}})$ , sendo que valores maiores que 1 indicam superioridade do modelo G-HO em relação ao T-HO.

## Resultados e Discussão

### Escolha do modelo

Baseado nos valores da função LVM para os modelos G-HO, T-HO, S-HO, G-HE, T-HE e S-HE, respectivamente, de -100.563, -99.623, -99.711, -99.455, -99.052, e -99.098, conclui-se que o modelo de resíduos com distribuição  $t$  de Student heteroscedástica forneceu o melhor ajuste aos dados de GPD entre as seis especificações diferentes consideradas. Este modelo teve um PBF  $9.328 \times 10^{19}$  comparado ao segundo modelo de melhor ajuste, o modelo Slash heteroscedástico.

Além disso, pode-se concluir definitivamente que o modelo Gaussiano homoscedástico, tipicamente utilizado nas avaliações genéticas, é a pior escolha para este conjunto de dados, uma vez que os valores de PBF desse modelo comparado com os outros cinco utilizados sempre se aproximaram de zero.

Pode-se, também, perceber pelos valores de LVM que a especificação de heteroscedasticidade foi relativamente mais importante que as caudas longas para as distribuições residuais, particularmente porque o modelo G-HE ajusta melhor os dados do que os modelos robustos homoscedásticos (T-HO e S-HO). Não obstante, a distribuição  $t$  de Student sempre ajustou melhor os dados comparados à distribuição Slash, quando as comparações foram feitas dentro das especificações de erro homoscedástico ou heteroscedástico. Portanto, as comparações dos demais parâmetros de interesse foram focadas na distribuição  $t$  de Student (melhor modelo robusto) e Gaussiana (modelo usual).

## Avaliação das Fontes de Heteroscedasticidade

A densidade a posteriori do efeito de sexo ( $\tau_1$ ) para a variabilidade residual foi muito similar nos dois modelos de erro heteroscedástico e, contrariando outros estudos para características de crescimento em gado de corte (GARRICK et al., 1989; RODRIGUES-ALMEIDA et al., 1995), sexo não foi uma fonte muito significativa de heteroscedasticidade residual (Tabela 1). Condições ambientais relativamente adversas podem ter contribuído neste sentido, tal que aos terneiros machos não foi permitido expressar completamente seu potencial extra de crescimento em relação às fêmeas, pois o ganho médio diário dos terneiros machos foi  $0,43 \pm 0,17$  kg e de  $0,34 \pm 0,13$  kg para fêmeas.

Não houve nenhuma evidência de que a variabilidade residual dependa da proporção racial já que o intervalo de probabilidade posterior de 95% (IPP) para  $\gamma_1$  inclui 1, o valor neutro em um modelo multiplicativo (Tabela 1). Já o efeito de heterozigosidade ( $\gamma_2$ ) na heteroscedasticidade residual foi mais convincente com a  $\Pr(\gamma_2 > 1 | y) = 0.0449$  para o modelo T-HE indicando que a medida que a heterozigosidade (e consequentemente, a heterose) aumenta, a variabilidade residual diminui. Este resultado é consistente com a teoria de que a heterozigosidade age como um tampão homeostático para minimizar a variação ambiental (LERNER, 1954).

A inferência a posteriori sobre o coeficiente de variação entre GC,  $\sigma_v$ , indica evidência de heteroscedasticidade residual significativa. É interessante notar que a média a *posteriori*  $\pm$  desvio-padrão a *posteriori* de  $\sigma_v$  ( $0,84 \pm 0,07$ ) sob o modelo Gaussiano heteroscedástico foi ligeiramente maior do que a estimada pelo modelo de  $t$  de Student heteroscedástico ( $0,72 \pm 0,06$ ). Este resultado decorre de a especificação de erro  $t$  de Student atenuar os efeitos de dados com resíduos extremos que tendem a inflacionar as estimativas da variância e consequentemente estimativas de  $\sigma_v$  sob uma especificação de erro Gaussiana.

**Tabela 1.** Média posterior (MEDIAP), desvio padrão posterior (DPP), moda posterior (MODAP), intervalo de probabilidade posterior de 95% (95% IPP), e tamanho efetivo de amostra (TEA) para os parâmetros de heteroscedasticidade residual no ganho pós-desmama obtidos pelos modelos de resíduos com distribuição heteroscedástica Gaussiana (G-HE) e  $t$  de Student (T-HE).

Parâmetro de variância <sup>a</sup>	MEDIA	DPP	MODAP	95 % IPP	TEA
<i>G-HE</i>					
$\sigma_e^2$ , kg <sup>2</sup>	365,85	15,43	393,86	[366,71, 427,42]	1,463
$\tau_1$	1,14	0,09	1,13	[0,98, 1,32]	1,712
$\gamma_1$	0,94	0,35	0,75	[0,43, 1,73]	146
$\gamma_2$	0,77	0,15	0,71	[0,52, 1,10]	142
$\sigma_v$	0,84	0,07	0,82	[0,71, 1,00]	1,142
<i>T-HE</i>					
$\sigma_e^2$ , kg <sup>2</sup>	388,75	15,27	387,37	[360,04, 419,59]	1,432
$\sigma_e^2$ , kg <sup>2</sup>	282,19	12,49	281,74	[258,41, 307,42]	859
$\tau_1$	1,13	0,09	1,11	[0,97, 1,31]	1,218
$\gamma_1$	1,15	0,45	0,80	[0,48, 2,20]	160
$\gamma_2$	0,70	0,16	0,61	[0,46, 1,06]	111
$\sigma_v$	0,72	0,06	0,71	[0,62, 0,86]	1,093

<sup>a</sup>  $\sigma_e^2$  = variância residual de referência,  $\tau_1$  = efeito multiplicativo de sexo (macho em relação a fêmea),  $\gamma_1$  = efeito da proporção de Nelore,  $\gamma_2$  = efeito da heterozigotidade,  $\sigma_v$  = coeficiente de variação para os efeitos aleatórios de grupos de contemporâneos, e  $\sigma_e^2 = \frac{v\sigma_e^2}{v_e - 2}$  = variância residual marginal para o modelo  $t$  de Student.

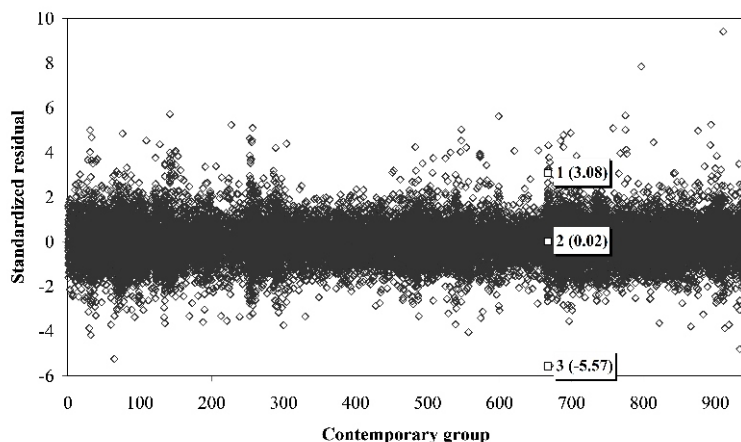
## Detecção e mitigação de dados extremos (“outliers”)

O gráfico de dispersão dos resíduos padronizados por GC, derivados da análise do GPD utilizando G-HO (Figura 1), apresenta 0,25% dos dados com resíduos (45 observações) fora do intervalo  $\pm 4,0$  erros padrão. Esta percentagem é 30 vezes maior à esperada quando é assumida distribuição normal homocedástica. Além disso, a curtose estimada desta distribuição residual padronizada foi de 2,72, indicando que a distribuição foi um tanto leptocúrtica. Estas características explicam os resultados de um ajuste substancialmente melhor dos modelos robustos em relação aos Gaussianos nesse conjunto de dados.

A densidade posterior de  $w_n$  nos modelos robustos prove informação útil para acessar o grau de afastamento de cada observação em relação a pressuposição de distribuição Gaussiana dos resíduos e, portanto, além de atenuar o efeito de observações extremas na estimação dos parâmetros, podem ser usados para detectar “outliers” (ROSA et al., 2003). Valores de  $w_n$  próximos a zero sugerem que a observação correspondente é extremamente afastada do seu valor predito pelo modelo, enquanto que valores de  $w_n$  próximos a 1 indicam que  $y_{ijklmn}$  é praticamente igual ao predito pelo modelo.

Para ilustrar este ponto num contexto de avaliação genética multirracial, foram selecionados deliberadamente para estudos posteriores observações de três machos  $F_1$  pertencentes ao mesmo GC de 160 animais da Região 2. Baseado nas especificações do G-HO, a primeira observação ( $y_1$ ) representa a um “outlier” moderado, estando 3,08 desvios padrão (DP) acima de zero, a segunda observação  $y_2$  representa um resíduo próximo a zero, ou seja, um ajuste perfeito do modelo (0,02 DP), e a terceira observação corresponde a um “outlier” extremo, estando 5,57 DP de zero (Figura 1).

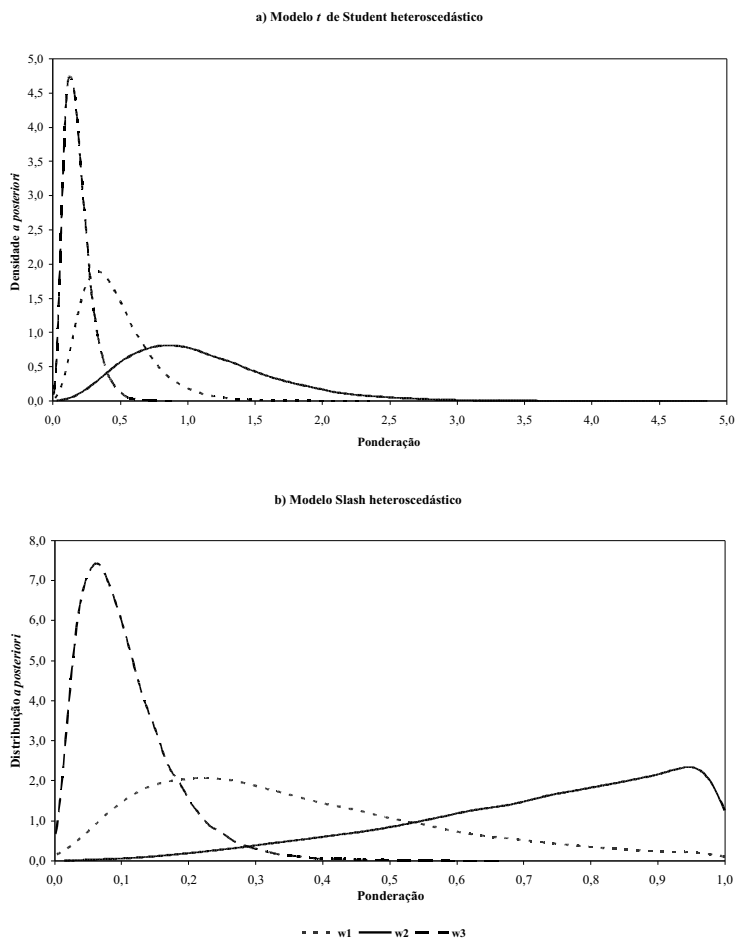




**Figura 1.** Gráfico de dispersão de resíduos padronizados para ganho pós-desmame por grupo contemporâneo utilizando modelo Gaussiano homoscedástico (G-HO). Três resíduos para o mesmo grupo contemporâneo são destacados para facilitar inferências que serão apresentadas posteriormente: 1. Representa um “outlier” médio +3,08 desvios padrão (DP) acima de zero; 2. Representa um resíduo próximo a zero (ajuste perfeito) e 3. Consiste em um “outlier” extremo, -5,57 DP de zero.

No caso de  $y_1$ , a média a posteriori de 0,46 e o IPP de 95% de [0,12; 1,07] para  $w_1$  com o modelo T-HE, sugerem que a influencia de  $y_1$  na inferência estatística é atenuada por este modelo, embora em menor extensão relativa a  $y_3$ . Esta baixa ponderação para  $y_1$  foi observada também com o modelo S-HE com uma média a posteriori e IPP de 95% para  $w_1$  de 0,36 e [0,06; 0,89], respectivamente.

Para a distribuição  $t$  de Student, a expectativa a priori para  $w_n$  é o valor neutro de 1,00; portanto, o fato do intervalo de confiança para  $w_n$  incluir ou não o valor 1,00 pode ser aplicado como critério objetivo para julgar se uma observação é um outlier (não incluir 1,00) ou não (inclui 1,00) em relação às especificações Gaussianas dos resíduos com um nível de precisão desejado. Por exemplo, o IPP de 95% pelo modelo T-HE contem 1,00 para  $w_1$  e  $w_2$ , mas não para  $w_3$ , indicando que somente  $y_3$  poderia ser considerado “outlier” por este critério.



**Figura 2.** Distribuição a posteriori das variáveis de ponderação correspondentes à observação 1 ( $w_1$  – “outlier” médio); observação 2 ( $w_2$  – modelo de ajuste quase perfeito) e observação 3 ( $w_3$  – “outlier” extremo) sob os modelos robustos: a) modelo  $t$  de Student heteroscedástico e b) Modelo Slash heteroscedástico.

Esse tipo de inferência posterior sobre  $w_n$  é uma estratégia que pode ser utilizada para monitoramento da qualidade dos dados. Por exemplo, se são observados baixos valores ( $< 1,00$ ) das variáveis de ponderação concentradas em determinados rebanhos ou se existe uma diferença substancial nas médias *a posteriori* dessas variáveis entre rebanhos. Nesses casos, rebanhos que têm alta frequência de médias *a posteriori* baixas para as variáveis de ponderação poderão ser identificados como tendo problemas na qualidade do processo de coleta de dados.

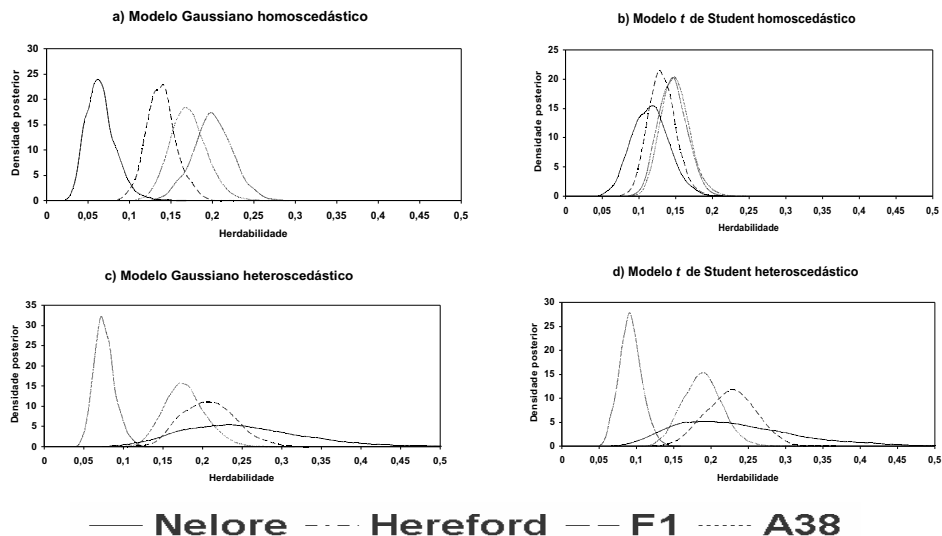
A edição de dados utilizada no melhoramento animal para determinar quais registros são “outliers” é frequentemente baseada em justificativas empíricas. Por exemplo, se a estratégia proposta por Bertrand e Wiggans (1998) da relação registro/média de subclasse fosse aplicada aos dados de GPD neste estudo, resultaria na eliminação de 1.515 registros (6,7% dos registros), os quais ficam fora do intervalo de 60% a 140%. Acreditando que somente uma parte destas observações corresponda a erros de registro, a aproximação de Bertrand e Wiggans (1998) parece ser um critério muito rígido para este conjunto particular de dados. Evidentemente, a edição e eliminação de registros devem ser implementados em situações onde erros de registros são óbvios e métodos para a detecção e correção de registros atípicos, como os utilizados no teste de produção diária de leite nos Estados Unidos (WIGGANS et al., 2003), fornecem proteções úteis para os sistemas atuais de avaliação genética.

Não obstante, é importante reconhecer que os modelos de caudas longas propostos ponderam cada observação diferencialmente para realizar inferências, de tal forma que os registros julgados como discrepantes ( $w_n < 1$ ) fornecem uma contribuição proporcionalmente menor na estimação de parâmetros quando comparados aos demais registros. Consequentemente, estes modelos fornecem um tratamento estatístico muito mais adequado aos registros extremos, que a alternativa de detectar e eliminar (ou ajustar) os “outliers” empiricamente.

## Herdabilidades

Apesar das mesmas especificações de variância-covariância multirracial (**G**) para os valores genéticos em **u** terem sido adotadas em todos os seis diferentes modelos empregados para analisar GPD, a inferência nos componentes de variância  $e$ , consequentemente, as herdabilidades mudaram consideravelmente dependendo da especificação para a distribuição dos resíduos.

As densidades a posteriori das herdabilidades para os modelos G-HO, T-HO, G-HE e T-HE são apresentadas na Figura 3. Sob os modelos de resíduos heteroscedásticos, as estimativas da herdabilidade para a raça Nelore foram maiores do que aquelas para a raça Hereford. Entretanto, o inverso foi observado sob os modelos de resíduos homoscedásticos. O grupo racial  $F_1$  tendeu a ter maiores estimativas de herdabilidade sob os dois modelos heteroscedásticos (G-HE e T-HE) comparados aos dois modelos homoscedásticos (G-HO e T-HO), em função da maior variância genética e menor variância fenotípica sob os modelos heteroscedásticos. As estimativas da herdabilidade nos quatro diferentes modelos foram similares para os animais enquadradas na composição racial do Braford (3/8 Nelore 5/8 Hereford), que tem somente 46,88% da heterozigosidade do  $F_1$ .



**Figura 3.** Densidades posteriores das herdabilidades para o ganho pós-desmama de grupos com diferentes composições raciais: Nelore, Hereford, F1 e Braford (3/8 Nelore 5/8 Hereford avançado - A38), obtidas por modelos com diferentes distribuições dos resíduos: a) Gaussiano homoscedástico, b)  $t$  de Student homoscedástico, c) Gaussiano heteroscedástico e d)  $t$  de Student heteroscedástico.

Os modelos heteroscedásticos melhor capturam a incerteza extra na estimativa dos componentes de variância para a raça Nelore. Isto é demonstrado na inferência da herdabilidade pelas distribuições posteriores mais planas cobrindo uma amplitude maior de valores possíveis (painéis c e d na Figura 3) e se deve, provavelmente, ao fato de que os animais da raça Nelore não possuem registros próprios. Neste caso, toda informação para as estimativas dessa raça é derivada da sua progênie cruzada.

Uma provável razão para a mudança substancial dos componentes de variância e da herdabilidade das raças Nelore e Hereford entre os modelos homoscedásticos e heteroscedásticos é que a maioria dos GC com grandes variâncias residuais (p. ex.,  $v_l > 3$ ) são compostos exclusivamente de animais da raça Hereford. Consequentemente, as estimativas da variância genética do Hereford podem estar viesadas para cima quando a variância residual é assumida como homoscedástica entre grupos raciais nos modelos G-HO e T-HO.

## Avaliações genéticas

Inferências sobre efeitos genéticos aditivos aleatórios são utilizados para ordenar animais para seleção e prever a diferença esperada na progênie (DEP) em raças puras. No entanto, em populações multirraciais, efeitos fixos aditivos e não aditivos assim como efeitos aleatórios não aditivos são de interesse, pois também determinam a diferença esperada na progênie multirracial (DEP-MR) (ELZO; WAKEMAN, 1998; POLLAK; QUAAS, 1998).

## Comparação dentro de composições raciais

As correlações de ordem de Kendall entre os modelos G-HO e T-HO, G-HO e G-HE, G-HE e T-HE e entre T-HO e T-HE para as médias a posteriori dos efeitos aditivos para os genótipos mais prevalentes são apresentados na Tabela 2.

**Tabela 2.** Correlações de ordem Kendall entre médias a posteriori de efeitos genéticos aditivos aleatórios para ganho pós-desmame para diferentes combinações dos modelos Gaussiano homoscedástico (G-HO),  $t$  de Student homoscedástico (T-HO), Gaussiano heteroscedástico (G-HE) e  $t$  de Student heteroscedástico (T-HE) para todos os animais e para o percentil 10% superiores para G-HO entre os genótipos mais frequentes.

Composição Racial	N <sup>1</sup>	G-HO vs. T-HO	G-HO vs. G-HE	G-HE vs. T-HE	T-HO vs. T-HE
<i>Incluindo todos os animais dentro de genótipo</i>					
Nelore	7.445	0,91	0,86	0,95	0,90
Hereford	19.976	0,84	0,76	0,89	0,84
F <sup>1</sup>	8.718	0,83	0,68	0,93	0,77
Braford <sup>2</sup>	1.452	0,85	0,82	0,92	0,88
<i>Considerando somente os animais classificados para os 10% superiores pelo modelo G-HO dentro de genótipos</i>					
Nelore	745	0,29	0,39	0,72	0,38
Hereford	1.998	0,36	0,31	0,77	0,67
F <sup>1</sup>	872	0,57	0,44	0,85	0,64
Braford <sup>2</sup>	145	0,37	0,28	0,63	0,55

<sup>1</sup> N é o tamanho de amostra para cada grupo de composição racial

<sup>2</sup> Composição racial 3/8 Nelore e 5/8 Hereford

A correlação de ordem entre os modelos do erro Gaussianos e  $t$  de Student dentro de qualquer classificação de heteroscedasticidade (G-HO vs. T-HO e G-HE vs. T-HE) foram razoavelmente altas entre os genótipos mais frequentes, sendo sempre superior a 0,83 para os modelos homoscedásticos e superiores a 0,89 para os modelos heteroscedásticos. Entretanto, quando se consideraram animais classificados como 10% superiores utilizando o modelo G-HO, a correlação de ordem entre valores genéticos aditivos para estes animais superiores pelos modelos G-HO e T-HO diminuiu consideravelmente (Tabela 2). Estes últimos resultados têm maiores implicações para as avaliações genéticas e as decisões de seleção que as correlações resultantes quando são envolvidos todos os animais. Essas correlações mais baixas para os animais superiores devem-se, ao menos em parte, ao fato que os registros discrepantes conduzem a predições mais extremas do mérito genético do animal no modelo G-HO, comparado ao modelo T-HO que tem poder de mitigação de dados extremos.

Isto é evidente na Figura 4 (gráfico superior), onde as médias *a posteriori* dos efeitos genéticos aditivos obtidos pelo modelo G-HO para Nelore, Hereford e

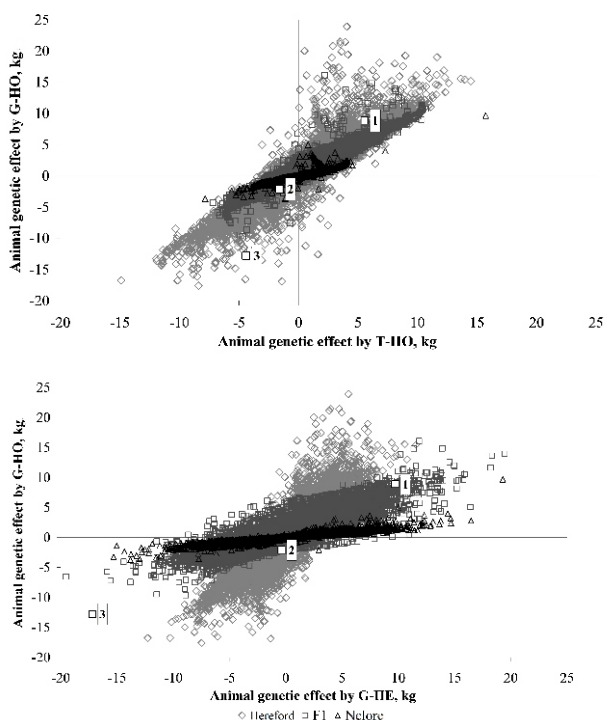
indivíduos  $F_1$  são apresentado frente às correspondentes médias a *posteriori* obtidas pelo modelo T-HO. Neste gráfico se observa que em alguns animais com médias a *posteriori* extremas para efeitos genéticos no modelo G-HO, tiveram estimativas mais próximas a zero no modelo T-HO. Por exemplo, o animal  $F_1$  associado com  $y_3$  (o “outlier” extremo descrito previamente) tem uma média a posteriori de -13 kg no modelo G-HO, mas somente -4 kg no modelo do erro T-HO (Figura 4 – gráfico superior). Estes resultados concordam com os resultados simulados de Straden e Gianola (1998) para modelo homoscedástico, onde o modelo  $t$  de Student atenua os efeitos de tratamentos preferenciais e conduz a inferências mais confiáveis dos efeitos genéticos associados com registros discrepantes. Foi observada também uma correlação de ordem mais baixa entre os modelos G-HE e T-HE quando se considera somente os animais 10% superiores que quando são considerados todos os animais (Tabela 2). Entretanto, o valor da redução na correlação e o grau de mudança na predição dos efeitos genéticos (Figura 5 – gráfico superior) não foram tão substanciais como entre os modelos do erro G-HO e T-HO. Os resultados sugerem que a modelagem de heteroscedasticidade residual é potencialmente mais importante que a modelagem de robustez para avaliações genéticas. Isto pode ser visualizado adicionalmente pela baixa correlação de ordem entre os modelos G-HO e G-HE comparadas com a dos modelos do erro G-HO e T-HO (Tabela 2).

## Comparação entre Composições Raciais

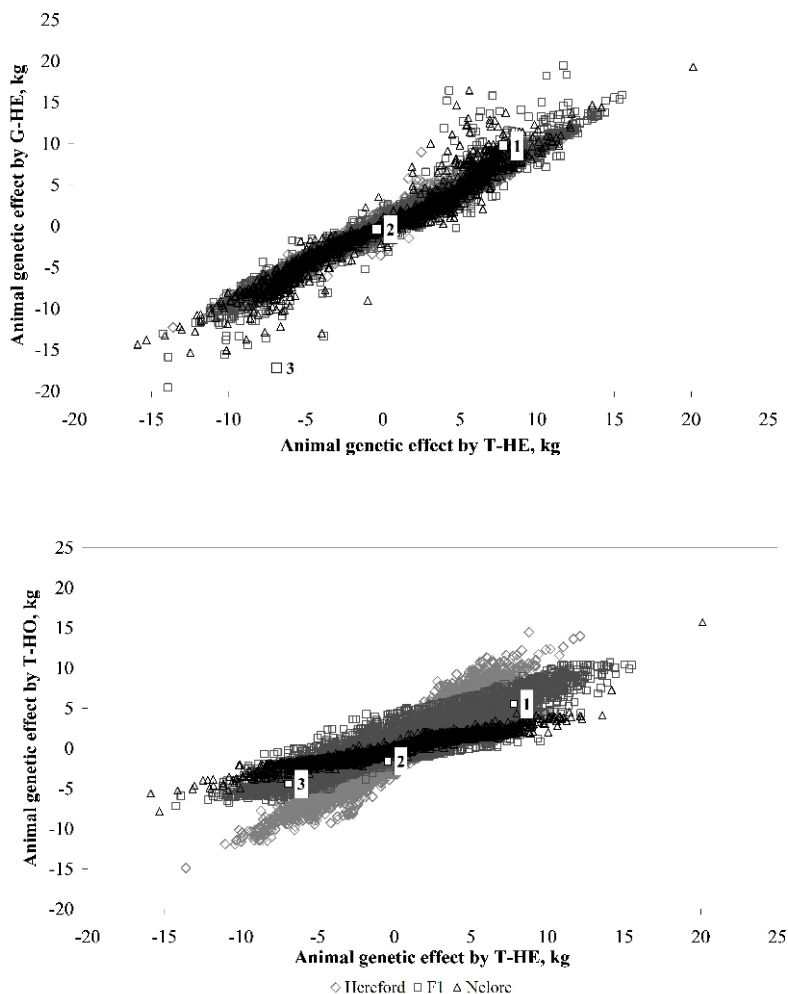
Se as decisões de seleção serão tomadas entre reprodutores de composições raciais diferentes, efeitos genéticos fixos aditivos ( $A$ ) e não aditivos ( $D$  e  $AA$ ) e a interação de  $A$  com outros efeitos como apresentado na Equação [1] devem ser considerados conjuntamente com os efeitos genéticos aleatórios aditivos para calcular diferenças esperadas na progênie multirraciais (DEP-MR). Neste caso, o desempenho esperado da progênie e suas diferenças serão função de coeficientes genéticos (composição racial, heterozigossidade e perdas por recombinação) que variaram de acordo com o genótipo do reprodutor e do seu par no acasalamento. Em outras palavras, a seleção do candidato entre composições raciais e a predição das DEP-MR devem ocorrer para um genótipo específico de acasalamento.



Para ilustrar este ponto, foram simulados acasalamentos de dois reprodutores  $F_1$  ( $F_1$ -1 com valores genéticos aditivos aleatórios de  $u = 7.8$  kg e  $F_1$ -2 de  $u = -0.4$  kg), que são os animais previamente considerados neste trabalho em relação às observações  $y_1$  e  $y_2$ , junto com dois animais Hereford e dois Nelore com efeitos genéticos aditivos idênticos. Foram comparados os genótipos da progênie futura, os coeficientes genéticos, as médias genotípicas, DEP-MR, o percentual de classificação das DEP-MR em relação a toda população, e sua classificação entre estes seis candidatos a seleção, de acordo com os genótipos das vacas foram acasalados (Tabela 3).



**Figura 4.** Médias posteriores de efeitos genéticos aditivos para ganho pós-desmame obtidos pelos modelos Gaussiano homoscedástico (G-HO) e t de Student homoscedástico (T-HO) (superior) e pelos modelos G-HO e Gaussiano heteroscedástico (G-HE) (inferior), para as composições raciais Nelore, Hereford e F1. Em destaque animais F1 pertencentes ao mesmo grupo contemporâneo e associados com registros representando no modelo G-HO: 1. “outlier” médio, sendo +3,08 desvios padrão (DP) de zero; 2. Resíduo próximo a zero (ajuste perfeito), e 3. “outlier” extremo, -5,57 DP de zero.



**Figura 5.** Médias posteriores de efeitos genéticos aditivos para ganho médio pós-desmame obtidos pelos modelos Gaussiano heteroscedástico (G-HE) e  $t$  de Student heteroscedástico (T-HE) (superior) e pelos modelos  $t$  de Student homoscedástico (T-HO) e T-HE (inferior), para grupos de composições raciais Nelore, Hereford e F<sub>1</sub>. Em destaque animais F<sub>1</sub> pertencentes ao mesmo grupo contemporâneo e associados com registros representando no modelo G-HO: 1. “outlier” médio, sendo +3,08 desvios padrão (DP) de zero; 2. Resíduo próximo a zero (ajuste perfeito), e 3. “outlier” extremo, -5,57 DP de zero.

**Tabela 3.** Genótipo da progênie e coeficientes genéticos para proporção Nelore (f), heterozigiosidade ( $\delta$ ) e perdas por recombinação (r), estimativas<sup>1</sup> de médias genótípicas da progênie, efeitos genéticos aditivos aleatórios (u), diferenças esperadas na progênie multirraciais (DEP-MR), percentis de DEP-MR frente à população total (%) e classificação relativa entre seis candidatos à seleção de diferente composição racial - dois Hereford (H), dois  $F_1$  e dois Nelore (N) – de acordo ao genótipo da vaca, para ganho pós-desmame obtidos pelo modelo  $t$  de Student heteroscedástico.

Candidato à seleção	Genótipo	Progenie			Média, Kg	u, Kg	DEP-MR, Kg (%)	Ordem
		Coeficientes genéticos						
		<i>f</i>	<i>δ</i>	<i>r</i>				
<i>Genótipo da vaca = Nelore</i>								
Hereford-1	F <sub>1</sub> ( HxN )	0,500	1,000	0,500	117,6	7,8	18,1 ( < 1 )	1
Hereford-2	F <sub>1</sub> ( HxN )	0,500	1,000	0,500	117,6	-0,4	9,9 (30)	2
F <sub>1</sub> -1	¾N ¼H	0,750	0,500	0,375	100,2	7,8	0,7 (54)	3
F <sub>1</sub> -2	¾N ¼H	0,750	0,500	0,375	100,2	-0,4	-7,5 (70)	4
Nelore-1	Nelore	1,000	0,000	0,000	88,9	7,8	-10,6 (76)	5
Nelore-2	Nelore	1,000	0,000	0,000	88,9	-0,4	-18,8 (91)	6
<i>Genótipo da vaca = Hereford</i>								
Hereford-1	Hereford	0,000	0,000	0,000	106,1	7,8	4,3 (22)	4
Hereford-2	Hereford	0,000	0,000	0,000	106,1	-0,4	-4,0 (73)	6
F <sub>1</sub> -1	¼N ¾H	0,250	0,500	0,375	109,6	7,8	7,7 (14)	3
F <sub>1</sub> -2	¼N ¾H	0,250	0,500	0,375	109,6	-0,4	-0,5 (40)	5
Nelore-1	F <sub>1</sub> ( HxN )	0,500	1,000	0,500	119,1	7,8	17,2 ( < 1 )	1
Nelore-2	F <sub>1</sub> ( HxN )	0,500	1,000	0,500	119,1	-0,4	9,0 (11)	2
<i>Genótipo da vaca = F</i>								
Hereford-1	¼N ¾H	0,250	0,500	0,375	108,8	7,8	11,2 ( < 1 )	1
Hereford-2	¼N ¾H	0,250	0,500	0,375	108,8	-0,4	2,9 (33)	4
F <sub>1</sub> -1	F <sub>2</sub>	0,500	0,500	0,500	101,8	7,8	4,2 (20)	2
F <sub>1</sub> -2	F <sub>2</sub>	0,500	0,500	0,500	101,8	-0,4	-4,0 (77)	5
Nelore-1	¾N ¼H	0,750	0,500	0,375	100,9	7,8	3,3 (30)	3
Nelore-2	¾N ¼H	0,750	0,500	0,375	100,9	-0,4	-4,9 (81)	6

<sup>1</sup> Estimativas relativas à média marginal posterior

Quando os animais Hereford-1, F1-1 e Nelore-1 são acasalados com vacas Nelore, apresentam DEP-MR substancialmente diferentes (Tabela 3), apesar de terem os mesmos efeitos genético aditivos aleatórios, sendo classificados em toda a população multirracial dentro do 1°, 54° e 76° percentis, respectivamente. Neste caso, as diferenças na DEP-MR são devidas as diferentes médias dos genótipos da progênie (Tabela 3). Entretanto, se estes mesmos três animais forem acasalados com vacas Hereford, suas DEP-MR serão classificadas em forma oposta entre o 22°, 14°, e 1° percentis, respectivamente. Além disso, entre os seis animais considerados neste exemplo na Tabela 3, Nelore-1 é classificado 1°, 3° ou 5° dependendo se o genótipo da vaca for Hereford, F1 ou Nelore, respectivamente. Estes resultados demonstram a forte dependência entre o genótipo do par no acasalamento e a classificação da DEP-MR entre as composições raciais. Finalmente, pode ser observado na Tabela 3 que a diferença de DEP-MR dos dois candidatos à seleção da mesma composição racial não muda através dos diferentes genótipos de acasalamento, sendo sempre igual à diferença entre seus efeitos genéticos aditivos; estes resultados dão suporte à afirmação de que a última informação basta para a seleção dentro de uma mesma composição racial, sempre que não sejam considerados os efeitos genéticos não aditivos aleatórios.

## Conclusões

A implementação de avaliações genéticas multirraciais robustas a dados extremos utilizando modelos baseados em variâncias genéticas e residuais heterogêneas é recomendada para se obter estimativas mais fidedignas e confiáveis do mérito genético de animais cruzados.

A característica de robustez do modelo permite ponderação diferente dos registros para inferência, tal que registros discrepantes têm contribuições menores às avaliações genéticas comparado aos registros mais plausíveis em consonância com as especificações do modelo, as características e os parâmetros genéticos da população.

Além disso, modelos robustos podem facilitar a identificação de sistemas de manejo e de grupos de contemporâneos com uniformidade superior do produto, um resultado cada vez mais importante na comercialização de bovinos de corte.

Especificações de heterogeneidade de variância melhor representam as distintas variabilidades genéticas e ambientais inerentes a populações multirraciais.

Neste estudo envolvendo cruzamentos entre as raças Nelore e Hereford, o modelo robusto heterogêneo t de Student apresentou o melhor ajuste aos dados e teve substanciais diferenças na classificação dos animais superiores para seleção comparado ao modelo normal homogêneo. Consequentemente, a utilização para avaliações genéticas de modelos baseados em resíduos com distribuição normal e com variâncias genéticas e residuais homogêneas podem diminuir o progresso genético em populações multirraciais de bovinos de corte.

## Referências

BERTRAND, J. K.; WIGGANS, G. R. Validation of data and review of results from genetic evaluation systems for us beef and dairy cattle. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 6., 1998, Armidale. **Proceedings...** Armidale: [s.n.], 1998. v. 27, p. 425-432.

CARDOSO, F. F.; ROSA, G. J. M.; TEMPELMAN, R. J. Multiple breed genetic inference using heavy-tailed structural models for heterogeneous residual variances. **Journal of Animal Science**, Savoy, v. 83, n. 8, p. 1766-1779, Aug. 2005.

CARDOSO, F. F.; TEMPELMAN, R. J. Hierarchical Bayes multiple-breed inference with an application to genetic evaluation of a Nelore-Hereford population. **Journal of Animal Science**, Savoy, v. 82, n. 6, p. 1589-1601, June 2004.

ELZO, M. A.; WAKEMAN, D. L. Covariance components and prediction for additive and nonadditive preweaning growth genetic effects in an Angus-Brahman multibreed herd. **Journal of Animal Science**, Savoy, v. 76, n. 5, p. 1290-1302, May 1998.

FOULLEY, J. L.; CRISTOBAL, M. S.; GIANOLA, D.; IM, S. Marginal likelihood and Bayesian approaches to the analysis of heterogeneous residual variances in mixed linear Gaussian models. **Computational Statistics and Data Analysis**, Amsterdam, v. 13, n. 3, p. 291-305, Apr. 1992.

GARRICK, D. J.; POLLAK, E. J.; QUAAS, R. L.; VANVLECK, L. D. Variance heterogeneity in direct and maternal weight traits by sex and percent purebred for Simmental-sired calves. **Journal of Animal Science**, Savoy, v. 67, n. 10, p. 2515-2528, Oct. 1989.

KINGHORN, B. The expression of recombination loss in quantitative traits. **Journal of Animal Breeding and Genetics**, Hamburg, v. 97, n. 2, p. 138-143, 1980.

KIZILKAYA, K.; TEMPELMAN, R. J. A general approach to mixed effects modeling of residual variances in generalized linear mixed models. **Genetics Selection Evolution**, Les Ulis, v. 37, n. 1, p. 31-56, Jan./Feb. 2005.

LANGE, K. L.; SINSHEIMER, J. S. Normal/independent distributions and their applications in robust regression. **Journal of Computational and Graphical Statistics**, Alexandria, v. 2, n. 2, p. 175-198, June 1993.

LERNER, I. M. **Genetic homeostasis**. Edinburgh: Oliver and Boyd, 1954. 134 p.

LO, L. L.; FERNANDO, R. L.; GROSSMAN, M. Covariance between relatives in multibreed populations - additive-model. **Theoretical and Applied Genetics**, New York, v. 87, n. 4, p. 423-430, Dec. 1993.

POLLAK, E. J.; QUAAS, R. L. Multibreed genetic evaluations of beef cattle. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 6., 1998, Armidale. **Proceedings...** Armidale: [s.n.], 1998. v. 23, p. 81-88.

RODRIGUEZ-ALMEIDA, F. A.; VANVLECK, L. D.; CUNDIFF, L. V.; KACHMAN, S. D. Heterogeneity of variance by sire breed, sex, and dam breed in 200-day and 365-day weights of beef-cattle from a top cross experiment. **Journal of Animal Science**, Savoy, v. 73, n. 9, p. 2579-2588, Sept. 1995.

ROSA, G. J. M.; PADOVANI, C. R.; GIANOLA, D. Robust linear mixed models with normal/independent distributions and Bayesian MCMC implementation. **Biometrical Journal**, Berlim, v. 45, n. 5, p. 573-590, 2003.

STRANDEN, I.; GIANOLA, D. Attenuating effects of preferential treatment with Student-t mixed linear models: a simulation study. **Genetics Selection Evolution**, Paris, v. 30, n. 6, p. 565-583, Nov./Dec. 1998.

WIGGANS, G. R.; VANRADEN, P. M.; PHILPOT, J. C. Detection and adjustment of abnormal test-day yields. **Journal of Dairy Science**, Savoy, v. 86, n. 8, p. 2721-2724, Aug. 2003.



---

*Pecuária Sul*

CGPE 9808